THE COST STRUCTURE OF REGIONAL TRANSMISSION ORGANIZATIONS

Daniel Greenfield

John Kwoka

Northeastern University

March 2010

Paper presented at the *US Energy Policy in Transition Conference*, Gainesville, FL. Gratitude is expressed to Jordi Jaumandreu, Kai-Uwe Kuhn, American Public Power Association, GDS Associates, and staff at the various RTOs for helpful comments and assistance with the data.

1. Introduction

For most of its history the U.S. electricity sector predominantly consisted of large, vertically integrated and heavily regulated utilities. By the late 1990s, however, a transformed industry took shape, characterized by substantial de-integration, a more competitive generation sector, and a greater market orientation to remaining regulation. These changes have required the development of several novel approaches and institutions to address unique features of electricity markets. Unlike most goods, electricity cannot be efficiently stored, and thus production must be constantly matched with final demand in order to avoid disruptions and possible blackouts. Within a vertically integrated firm, this coordination is achieved by common ownership and control of generating units and distribution systems. A system operator takes various system constraints into account and administratively ensures electrical balance at all times.

Whatever the merits of deintegration of generation for purposes of fostering competition in that market, deintegration creates the need for some alternative method of performing the coordination functions that were previously internal to the firm. All restructured electricity markets have had to develop some method of coordination. In the U.S. that method has been regional transmission organizations, or RTOs. RTOs are independent non-profit corporations that have functional control of transmission assets in a region, assets that are, however, owned by other entities, namely, the operating utilities that distribute electricity to end users. Most fundamentally, RTOs are intended to ensure equal access to the transmission network as they centrally dispatch electricity over member utilities' transmission systems. RTOs also organize competitive wholesale markets, coordinate the complex operating decisions of member utilities, monitor wholesale market performance, and coordinate long term investment planning.

RTOs now cover well over one-half of customers and sales of electricity in the U.S.[1] As they have expanded in geographic coverage and functional scope, controversies have arisen about their rapidly growing costs as well as their effectiveness in preventing congestion and in planning investment. This paper examines the first of these issues, namely, RTO costs. We exploit the fact that existing RTOs have come into being and initiated specific functions at different times. We relate those functions and their timing to differences in their costs over time in order to cast light on several questions:

First, we investigate the costs of each of the market functions administered by RTOs—day-ahead markets, financial transmission rights, and so forth. This will help in understanding which functions are most costly and thereby cast light on the relative cost of RTO functions.

Second, we investigate the relative costliness of individual RTOs in performing the same functions, in order to determine which represents the best (and worst)-practice RTO.

Third, we examine whether RTOs in operation for longer periods realize cost savings over time and also whether later-forming RTOs are less costly relative to early RTOs. Both of these effects represent types of learning economies.

Finally, we document economies of scale over the sampled range of RTO "output."

Not surprisingly, we find evidence that administering markets for energy, ancillary services, and financial transmission rights all increase RTO costs; however, we are unable to precisely estimate the contribution of each. In terms of differences in efficiency across RTOs, we find some evidence that ERCOT is the least costly. Finally, while we fail to find evidence of economies of learning, either within or across RTOs, we do document economies of scale.

[1] According to the ISO/RTO Council (IRC), an industry organization consisting of representatives of North American ISO/RTOs, RTOs serve more than 50 percent of Canada's population as well. See, http://www.isorto.org/.

The plan of this paper is as follows. The next section provides some background on wholesale electric power markets and the origin of RTOs. Section 3 further describes the evolution of existing RTOs and their various specific functions. Section 4 sets out the econometric model, hypotheses to be tested, and the data. Section 5 presents the results of our analysis and 6 concludes.

2. Background on regional transmission organizations

This section provides some history of RTOs, a description of their basic functions, and a discussion of controversy that has arisen regarding their costs.

2.1. *History*

For most of its history the U.S. electricity sector has been dominated by vertically integrated and heavily regulated utilities that bundled and sold generation, transmission, and distribution services. States were responsible for regulating retail electricity rates, and the Federal Energy Regulatory Commission (FERC, formerly the Federal Power Commission) was responsible for the rates, terms, and conditions of wholesale transactions and interstate transmission. Beginning in 1978, however, reforms made inroads on this traditional structure, and under the direction of Congress, FERC began creating opportunities for utilities and independent power producers to trade power at market based rates.[2] Some regions, particularly the Northeast, transformed their transmission systems from a set of balkanized utilities to

[2] The initial reforms of the industry spawned from the Public Utilities Regulatory Policies Act of 1978 (PURPA) and the Energy Policy Act of 1992 (EPACT 1992). Under these congressional directives, FERC developed opportunities for nonutility power producers and generation dependent utilities to sell and purchase wholesale power at market-based rates.

coordinated power pools aimed at facilitating wholesale transactions over larger geographic areas. Some of these power pools would eventually form RTOs.[3]

Initial efforts to foster wholesale competition were compromised by a transmission network optimized for vertically integrated utilities but also still largely owned by integrated utilities that were foreclosing access to their sections of the transmission network. Nonutility power producers and generation-dependent distribution utilities relied on the transmission wires owned by vertically integrated firms; thus, wholesale competition could not flourish without equal access to the grid. In 1996 FERC issued Order 888 requiring utilities to "functionally unbundle," that is, to state separate rates for wholesale generation, transmission, and ancillary services, as well as to take transmission service under the same tariff as it provided its transmission customers.[4]

While functional unbundling established a commitment to equal access, it was insufficient to overcome the inherent advantages of common ownership of transmission and distribution assets by most utilities. Accordingly, in December 1999 FERC issued Order 2000 with the goal that all transmission owners would place their transmission facilities under the control of an RTO.[5] FERC set minimum characteristics and functions for RTOs, namely, (1) independence from market participants, (2) possession of operational authority for all transmission facilities under the RTO's control, and (3) exclusive authority to maintain short-

[3] The RTOs in the Northeast developed from tight power pools, which functioned as one control area. Unlike RTOs, they did not have control of transmission facilities, nor were they independent from market participants.

[4] These tariffs set forth the price and non-price terms for providing transmission services, both energy and ancillary. Moreover, power pools and public utility holding companies were required to file a single tariff for all affiliated firms.

[5] While FERC has endorsed RTOs in Order 888, and some areas of the country were moving toward the RTO model, there was no previous requirement to establish such organizations.

term reliability. FERC argued that these institutions would bring about increased efficiency through regional transmission pricing; improved congestion management and grid reliability; more efficient planning for transmission and generation investments; more accurate estimates of available network capacity; and more effective management of physical flows of electricity. [6]

The seven existing RTOs are listed in Table 1. Despite diversity in implementation timelines and groupings of market functions, some broad trends emerge in recounting their evolution. The RTOs in the Northeast (New England ISO, or ISO-NE; New York ISO, or NYISO; and PJM[7]) share a regional resemblance. Each evolved from power pools and now operate a suite of wholesale markets—day-ahead and real-time markets for energy and ancillary services with nodal prices and financial transmission rights—that closely correspond to FERC's SMD proposal (see note 6). The Northeast RTOs were all operating by 2000, with ISO-NE the last to implement the SMD style market in 2003. Along with the Northeast RTOs, the California ISO (CAISO) and the Electric Reliability Council of Texas (ERCOT) were also early adopters of the RTO model; however, their approach to pricing in energy markets differed insofar as it relied on a single energy market with zonal pricing.

The Midwest ISO's (MISO) began in March, 1998, but its further development was characterized by lengthy negotiations among potential stakeholders. Substantive operations began in 2002 when MISO initiated its tariff administration function, however, it did not begin administering market functions until 2005. The Southwest Power Pool (SPP) has been the

[6] It is also worth noting that in 2002 FERC more aggressively sought to establish RTOs through a further rulemaking. The plan for a so-called Standardized Market Design (SMD), however, encountered fierce opposition and was abandoned.

[7] PJM serves Pennsylvania, New Jersey, and Maryland, as well as Delaware, Illinois, Indiana, Kentucky, Michigan, North Carolina, Ohio, Tennessee, Virginia, West Virginia and the District of Columbia.

slowest developing RTO. After incorporating in 1994 it began administering open-access transmission service on January 1, 1998 and implemented its first market function in 2007 when it launched its real-time energy market.

2.2. *Functions of RTOs*

RTOs are most unusual economic institutions. They are not-for-profit corporations that assume control and management of the bulk power systems of their member utilities, while the latter continue to own all of those assets. There are few entities with comparable ownership and control arrangements. RTOs are self-governed companies directed by a board that is independent from market participants but in many cases intended to reflect all stakeholders. RTOs raise capital and receive revenues for transmission services they provide. RTOs do not own transmission facilities, nor are they responsible for the maintenance and repair, or fixed investment costs, of the transmission facilities over which they direct the flow of power. Their essential role is as independent service provider that administers the terms and conditions of transmission services and maintains the short-term reliability of the network. As such, the organization is designed to ensure non-discriminatory access and operate the grid in a manner consistent with engineering efficiency and market competition.

RTOs perform a number of specific functions, each of which gives rise to added costs that will be examined below. Their descriptions are as follows:

2.2.1. Energy markets

RTOs create and operate decentralized trading markets for electrical energy. The most basic approach involves a real-time energy market that resolves discrepancies between real-time demand from load-serving entities and supply from producers. We will refer to this approach as the single-settlement energy market. This type of energy market has been employed largely by

ERCOT, CAISO, and SPP[8]. Other RTOs, including ISO-NE, NYISO, PJM, and MISO, have implemented two-settlement energy markets, where the RTO also operates a day-ahead forward market. The latter are intended to facilitate reliability and efficiency by allowing the RTO to secure adequate supplies to meet forecasted demand ahead of real-time, and to provide buyers and sellers the opportunity to hedge real-time price volatility.

2.2.2. Ancillary service markets

Since electricity cannot be effectively stored, system operators keep an inventory of generators available to call on when system conditions change. Examples of so-called ancillary services include spinning reserve, non-spinning reserve, reactive power, and regulation, each of which represents somewhat different degrees of contingent-available supply to the RTO. Ancillary services can be procured by the RTO via cost based payments, or less frequently, the RTO may organize a market where generators bid for the right to supply. For our purposes all of these variations will be considered as a single "ancillary services" market function.

2.2.3. Transmission Rights

To address possible network congestion, RTOs must have a system for allocating network capacity and keeping the system within its security limits. Some RTOs, including the Northeastern RTOs and MISO, employ "nodal pricing" in which participants submit bids to buy or sell electricity at specific network locations or "nodes," whereas other RTOs use a broader "zonal" pricing system. In addition, the nodal markets all use a so called security-constrained economic dispatch—a network-wide optimization process designed to choose the lowest-cost

[8] PJM and ISO-NE initially employed single-settlement energy markets.

bids to balance electricity supply and demand subject to the physical constraints of the system's generation and transmission facilities. All of these systems create price uncertainty, and so RTOs have created various forms of transmission rights to permit market participants to hedge the resulting uncertainty. Typically, the RTO organizes monthly and yearly auctions for transmission rights of various forms.

2.2.4. Resource adequacy and capacity markets

Although settlement markets were originally believed sufficient to guide market participants with respect to capacity investment, in practice this has not occurred. As a result RTOs have devised various supplementary incentives to foster investment. The Northeastern RTOs have all implemented markets for capacity in which demand side participants are required to procure a sufficient portfolio of generating capacity to meet a resource adequacy requirement. The RTO conducts auctions that can be used by participants that have not met these requirements with bilateral contracts. CAISO has also implemented a resource adequacy program; however it continues to evaluate whether a capacity market will be beneficial. Our empirical analysis seeks to determine the costs associated with capacity markets, but as we shall see the data make this part of the inquiry less satisfactory.

2.3. *Issues in RTO markets.*

In its original rulemaking on RTOs, FERC asserted that RTOs would be a cost effective method of achieving their important purposes. FERC staff estimated formation of RTOs would result in cost savings of \$2.4 billion annually.[9] While acknowledging nontrivial costs, FERC

[9] Order No. 2000, *Regional Transmission Organizations*, F.E.R.C. STATS. & REGS, p. 94.

predicted that later forming RTOs would experience learning economies from observing the initial forming RTOs. FERC explained that having the opportunity to observe "what works and what does not with respect to regional transmission entities . . . should make it somewhat easier, and more cost efficient, to create new RTOs." [10] Based on this assessment, it claimed that RTOs would confer considerably greater benefits than their costs.

As concerns about RTO costs emerged and grew more serious, FERC staff issued a report in 2004 that estimated the costs for developing and operating a "Day One RTO," specifically, an investment outlay of between $38 million and $117 million and an annual revenue requirement of between $35 million and $78 million. [11] A Day One RTO was described as satisfying the minimum requirements of Order No. 2000, including open access transmission service, scheduling authority, and certain other functions, but excluding security constrained economic dispatch, nodal prices, financial transmission rights, and other marker functions that were called Day Two functions. Since the latter were central features of four of the five RTOs operating at the time, FERC's study was sharply criticized as ignoring reality. The Public Power Council (PPC), for example, described FERC's analysis as static, entirely ignoring well documented escalating[12] costs, with its focus on Day One costs only serving to "confuse the discussion and obfuscate the true costs to customers."

---

[10] *Id.* at pp. 70-89.

[11] *Staff Report on Cost Ranges for the Development and Operation of a Day One Regional Transmission Organization*, FERC, October 2004.

[12] See, Lutzenhiser (2004) for a well documented account of escalating RTO costs, and a critical assessment of FERC's predictions in Order No. 2000, the Energy Information Administration's (2003) study of the *Impact of the Federal Regulatory Commission's Proposed Standard Market Design* and ICF Consulting's (2002) *Economic Assessment of RTO Policy*. Using updated cost data, she shows even optimistic projected benefits are outweighed by increasing RTO costs, concluding that FERC appears to be a "true-believer" in competitive markets and its RTO policy is not "based on a comprehensive analysis of the possible costs, benefits and risks."

Some private studies reinforced these concerns. Tracking past RTO costs and using proposed budgets, Lutzenhiser (2004) calculated that RTO expenditures would total approximately $1 billion. Extrapolating to the entire national electricity market, she estimated national RTO expenditures would be $2.4 billion, which was precisely FERC's own "best estimate" of the potential annual benefits of national RTO formation, excluding any development costs or indirect costs borne by market participants. Two consultancies, Christensen Associates (2007) and GDS Associates (2007), have also documented the significant operating expenditures of RTOs, noting that costs, in dollars per megawatt terms, vary widely across RTOs.[13]

In May of 2007, the U.S. Government Accountability Office (GAO) entered into this discussion. The GAO found a "general trend of economies of scale" and on a per megawatt hour basis ISO New England, Midwest ISO, and New York ISO expenses increased from 2002 to 2006, while CAISO, PJM, and SPP's expenses decreased. It recommended that FERC systematically review RTO budgets and annual financial reports and develop standardized measures to track RTO performance.[14] More critically, it found that FERC officials, industry participants, and experts disagreed on whether RTOs have brought benefits to their regions. Prominent concerns include swelling costs, a second-best governance structure, and an apparent inability of RTO run wholesale markets to stimulate adequate investment in transmission and generation capacity.

Two other concerns about RTOs deserve mention, although they are not part of our study. First, RTO governance has come under scrutiny. The RTO structure—a nonprofit that operates

[13] GDS predicts that as "RTOs expand their services and as their systems age and must be replaced, the administrative and operational costs will increase."

[14] See, Drom (2007).

assets they do not own—represents a compromise in the restructuring process, due to the existing balkanized transmission system. As Joskow (2005a) describes,

> Order 2000 effectively takes the existing ownership structure as a constraint and promotes the creation of new not-for-profit independent system operators (ISO, RTO, ITP, pick your favorite name) to deal with these issues. However, these independent entities own no transmission assets, have no linemen or helicopters to maintain transmission lines and respond to outages, and are not directly responsible for the costs of operating, investing in, or the ultimate performance of the transmission networks they 'manage.'

Ownership, maintenance, and investment decisions remain compartmentalized within individual utilities, while day-to-day system operations are managed by a central independent party. This bifurcation of transmission governance presents challenges for regulators, as well as complicating investment decisions for firms.[15]

Further, and perhaps most fundamentally, many observers feel that RTOs have simply failed to perform their intended functions well, at whatever cost.[16] Critics point to the increase in congestion, the lack of investment, and the overall performance of RTO-organized markets. There is broad consensus that transmission investment is failing to keep pace with growing demand for electricity, partially due to a failure in RTO regions to effectively develop a framework for regulating and allocating the cost of transmission investment. These criticisms have grown over time, bringing RTO expansion to a halt as policymakers search for an alternative model for performing the necessary functions.

Bearing these further controversies in mind, we now return to the issue of RTO costs.

---

[15] See, Joskow (2006) for a further discussion on the challenges in regulating transmission and distribution networks.

[16] For two broad ranging analyses of the benefits and cost of RTOs, which also include discussions of RTO operating expenditures, see: Seth Blumsack (2007) and Christensen Associated Energy Consulting (2007).

3. Empirical model

Our primary empirical model is a Cobb Douglass cost function that has been augmented to account for learning and the possibility that market functions may increase the amount of labor and capital needed to oversee transmission of a given amount of electricity. The underlying Cobb Douglass production function is $Q_{it} = \exp\,(c + \gamma' E_{it} - \delta' M_{it}) K_{it}^{\alpha} L_{it}^{\beta}$. Where c is an arbitrary constant, E_{it} measures experience both within and across RTOs, and M_{it} is a vector of indicator variables representing market functions. Capital and labor inputs are represented by, K_{it} and L_{it}, and output, Q_{it}, is annual net energy transmitted to load. There are many potential measures of the scale of an RTO, including the number of market participants, number of transmission owning utilities, and miles of transmission wire. We employ annual net energy transmitted to load due to its homogenous nature across RTOs, and straightforward interpretation. It should also be noted, this variable can be interpreted as proxy for other measures of RTO size.[17]

Once the RTO implements a set of market functions, we assume it minimizes cost subject to its production technology and the price of inputs. Additional market functions are hypothesized to increase costs as the RTO must acquire more capital and labor, while experience may reduce costs through economies of learning. The augmented Cobb Douglass cost function can be written:

$$(1) \quad \ln(TC) = \mu + \gamma' E_{it} + \delta' m_{it} + \frac{1}{\alpha+\beta} ln(Q_{it}) + \frac{\alpha}{\alpha+\beta} \ln(v_{it}) + \frac{\beta}{\alpha+\beta} \ln(w_{it}),$$

[17] We have experimented with more than one measure of output in the model. Unfortunately—likely due to already limited degrees of freedom—none of the other measures have yielded satisfactory results.

where the price of capital and labor are denoted by, v_{it} and w_{it}, respectively. Finally, we may write the empirical cost function as:

$$(2)\quad \ln(TC_{it}) = \mu + \gamma' E_{it} + \delta' M_{it} + \beta_1 ln(Q_{it}) + \beta_2 \ln(w_{it}) + \eta_i + \tau_t + \varepsilon_{it}.$$

η_i represents time-invariant effects specific to each RTO, τ_t is a term for yearly industry-wide effects, and ε_{it} represents an idiosyncratic error.[18] The η_i may include unobserved or difficult to quantify differences across RTOs, including differences in market monitoring, planning and expansion, and interregional coordination procedures. However, to the extent that unobserved cost causative functions are insignificant, our estimates of η_i, the RTO fixed effects, identify time-invariant differences in efficiency.

The temporal shocks, τ_t, include national changes in RTO policy and the price of capital.[19] We are particularly concerned about the effect of FERC rulemakings and policy statements that may alter RTO behavior or production technology. Table 4 lists and briefly describes the major FERC orders associated with RTOs. Controlling for time fixed effects is not only important for identifying the marginal impact of market functions, but is also critical for indentifying η_i. If the τ_t tend to be larger in later years, which we in fact observe, these positive

[18] In further work we intend to utilize the stochastic frontier model outlined by Greene (2005) "Fixed and Random Effects in Stochastic Frontier Models," *Journal of Productivity Analysis*. With this approach, the idiosyncratic component of the error will be further disaggregated as $\varepsilon_{it} = v_{it} + Su_{it}$, such that $v_{it} \sim \mathrm{N}[0, \sigma_v^2]$ captures random disturbances, and $u_{it} = |U_{it}|$ where $U_{it} \sim \mathrm{N}[0, \sigma_v^2]$ captures operating inefficiency.

[19] The price of capital may fluctuate with the interest rate, investor perceptions of the risk of financing an RTO, as well as changes in the price of underlying capital goods (primarily software).

effects will be correlated with the fixed effect terms of later forming RTOs, and omission of τ_t will tend to impart a positive bias on the fixed effects of later forming RTOs.

One of our primary goals is to estimate the incremental impact of market functions on RTO costs. This will be captured by δ, the coefficient vector on M_{it}. Given the diverse groupings and implementation timelines of the various market functions across RTOs, we have sufficient variation in these dummy variables to identify the relative costs of each market function. With the RTO and yearly fixed effects, with-in RTO variation in costs identifies the marginal effects of each market function. In other words, δ captures the difference-in-differences: the marginal effect of a particular market function on costs, after controlling for RTO-specific differences in costs (η_i) and industry-wide changes in costs (τ_t).

We also investigate economies of learning, that is, reductions in costs due to experience with the production process.[20] The coefficient vector γ captures the marginal effect of experience. We measure internal experience with a variable that counts the number of years of operation by an RTO, its age. Our measure of external experience is the total number of years of operation across all RTOS. That is, the sum of all RTOs' ages in a particular year. The coefficient on this variable is intended to capture the impact of industry-wide learning.

Finally, we note that the model estimates parameters of the Cobb Douglass production technology. The estimated elasticity of cost with respect to the wage is captured by β_2, and returns to scale can be measured as $1/\beta_1$, the reciprocal of the elasticity of total costs with respect to output. As one can see, the price of capital has been excluded from the empirical model. Unfortunately, the majority of an RTO's capital investments are in the development of

[20] Relevant studies of learning economies include Thornton and Thompson (2001), Gruber (1998), and Irwin and Klenow (1994).

proprietary software systems used to operate wholesale markets. As a result, we were unable to construct a relevant variable for the price of capital.[21]

We should also note possible concerns about endogeneity of RTO formation and its effect on our empirical results. Casual inspection of trade publications and FERC rulemakings indicates each RTO's initial market design arose from negotiations among stakeholders, the region's historical industry structure, and possibly some attention to performance.[22] For example, the RTOs in the Northeast share a common design that many observers attribute to a history of power pooling among utilities. Nevertheless, we believe that changes in market design throughout our sample period have been largely the product of FERC rulemaking and other mandates exogenous to any particular RTO, rather than the operating efficiency of the RTOs themselves.

4. Data and econometric issues

Our analysis is based on an unbalanced panel of annual data for the RTOs operating within the United States from 1998 to 2008. All but ERCOT, which operates solely within the borders of Texas, are subject to FERC jurisdiction and must submit annual financial statistics via the Form 1 electric utility annual report. These Form 1 data serve as our primary source for

[21] Using the generic property that a cost function is linearly homogenous in factor prices, the implied elasticity of total costs with respect to the price of capital is $1-\beta_2$.

[22] Paul Joskow (2005b) articulates a frequently echoed sentiment describing the process leading to California's initial market design as, "contentious and highly politicized, reflecting perceptions by various interest groups about how different wholesale market institutions would advance or constrain their interests and, in my view, an inadequate amount of humility regarding uncertainties about the performance attributes of different institutional arrangements. The discussion of alternative institutions was polluted by an unfortunate overtone of ideological rhetoric that attempted to characterize the debate about wholesale market institutions as one between 'central planners' and 'free market' advocates."

annual operating expenses. Data for ERCOT and for Southwest Power Pool (prior to being officially recognized by FERC in 2004) were collected from publicly available annual reports.[23]

The variable *OPEXP* represents the natural log of the standard annual accounting measure "total operating expenses" including operation expenses, maintenance expenses, and depreciation expenses. *LOAD* measures output, the log of annual net energy for load (GWh). The relevant data were compiled from RTO and FERC sources. Our series for the log of wage, *WAGE*, was constructed from Bureau of Labor Statistics data on the average annual pay of workers in the electric power generation, transmission, and distribution sector (NAICS 2211) of the state where the RTO's headquarters is located. Both wage and operating expense data are deflated to year-2005 dollars with the GDP deflator.

The dummy variables that describe the grouping of markets operated by RTO *i* in year *t* were constructed based on information from RTO websites, particularly the "state of the markets" reports prepared by each RTO's independent market monitor. The dummy variables assume a value of unity in any year in which the RTO implemented the specified market function(s) prior to the fourth quarter.[24] Table 2 describes the implementation timeline of market functions for each RTO. With the information in Table 2 in mind, we can explain how the market function dummies are defined. The first variable, *REALTIME*, indicates operation of a real-time energy market. The second variable, *DAYAHEAD_FTR*, equals one if RTO *i* operated a day ahead market with financial transmission rights in year *t*. We have combined the indicator

[23] We also experimented with inclusion of data from the RTOs operating in Canada. Unfortunately we found these institutions were likely described by meaningfully different production technologies, potentially due their composition of fewer transmission owners or different geography of their bulk transmission grids.

[24] PJM implemented a regulation market in June of 2000 and a market for reserves in December of 2002. As a compromise, we set the *ANCILL* equal to one for all years after 2000 in PJM.

variables for day-ahead energy and financial transmission rights because these market functions were simultaneously implemented by ISO-NE, MISO, NYISO, and PJM (implemented in the same 11 month period by PJM). *DAYAHEAD_FTR*, combined with *REALTIME*, represents the two-settlement energy markets with nodal pricing employed in the Northeast and Midwest. Because CAISO and ERCOT—the two RTOs that utilized a zonal market structure—did not simultaneously implement financial transmission rights with their energy markets, we are able to identify the impact of zonal financial transmission rights on RTO costs with the dummy variable *ZONAL_FTR*. Finally, *ANCILL* and *CAPACITY*, indicate operation of ancillary service and capacity market functions.

We recognize the use of a binary variable may well exaggerate how discrete the policy change actually is. Certainly, RTOs perform research and testing prior to market implementation, and cost effects may be greater in the first year of implementation. We view our current approach as a first approximation to be refined based on further information and testing.

Because of the limited size of our dataset, we face a delicate trade-off between accuracy and precision. Failing to control for the RTO-specific and year-specific components of the error leaves our estimates subject to omitted variables bias. However, due to the limited sample size and high correlation among many independent variables, including a full set of RTO and yearly fixed effects limits the efficiency of our estimates. Accordingly, to offer some initial descriptive results, we begin with several parsimonious specifications that do not fully account for unobservable factors. Then we sequentially add controls to more accurately identify the impact of market functions on costs, albeit at the cost of precision reflected in the standard errors. Finally, we augment the model to test for learning economies. This approach allows us to develop a set of results from which we can address potential sources of bias, and appropriately temper our conclusions given the limitations of the dataset.

To correct for group-wise heteroskedasticity and first order autocorrelation, we fit all specifications using a feasible GLS estimator. The estimated variance matrix is constructed from the residuals of an OLS regression, and allows for a unique variance and autocorrelation parameter in each panel.

5. Results and Discussion

This section begins with an overview of the results. We also outline our approach of sequentially adding additional controls. In subsections 5.2 – 5.5 we respectively discuss in more detail the cost effects of RTO market functions, the relative efficiency of particular RTOs, economies of learning, and economies of scale.

5.1. *Overview of Results*

The first column of table 5 presents the estimates of a model that does not control for industry-wide time-varying effects. While a causal interpretation of this model requires the assumption that τ_t is orthogonal to the included regressors—a very strong and likely unrealistic assumption—this model is included as a descriptive baseline. In column 2, a linear trend variable is added to capture time-varying unobserved factors. These estimates are more reasonable in terms of the magnitude and sign of the coefficients on the market function variables and *WAGE*. Even so, while the linear trend variable is quite intuitive, it is a very restrictive way to capture time-varying omitted factors.[25] Accordingly, column 3 presents the fully specified model with RTO and yearly fixed effects. With this specification, the model indicates that all functions, except administering a capacity market, add to costs.

The models presented in columns 4 through 6 explore economies of learning and scale. By adding the variable *LNYRSOP*, the first specification examines economies of internal learning, that is, efficiency gains associated with an RTO's own experience. Specification 5 adds *LNYRSOPTOTAL* to test for external learning arising from total industry experience. The parameter estimates of both models are not consistent with learning economies. In the final column, PJM and MISO have been excluded because of their geographic expansions over the sampled time period. When PJM and MISO are excluded, the coefficient on *LOAD* has a more obvious interpretation; that is, the elasticity of total cost with respect to load growth over a fixed network. That the coefficient on *LOAD* increases, suggests that some of the "economies of scale" observed with a full sample arise from geographic expansion.

5.2. *Market Functions*

We interpret the estimated cost effect of individual market functions with some caution, due to the aforementioned limitations of the market function dummies and small data set. Scanning across the models presented in table 5, one can see the parameter estimates are sensitive to the model specification. Nonetheless, in the most comprehensive specification presented in column 5, the estimated coefficients on all of the market function dummies absent

[25] There is no particular reason why temporal shocks should increase (or decrease) in a constant linear fashion. For example, the price of capital may fluctuate up-and-down with the interest rate, lenders perceptions of risk, and changes in the price of high technology capital goods.

CACPACITY have the expected sign and pass a one-sided test of statistical significance at the ten percent level (there is no hypothesis that market functions reduce cost).

The negative estimated coefficients on *CAPACITY* in the first two columns of table 5 are likely a result of omitted variables bias. Nonetheless, the non-effect estimated in the full model, is somewhat surprising. One possible explanation is that capacity markets have a trivial impact on costs, particularly compared to markets for energy or ancillary services. Even so, limitations of the data should be noted. Capacity markets have only been implemented by the Northeastern RTOs, and *CAPACITY* is highly correlated with other market function dummies.

One interesting policy implication from our results is that it appears the single settlement zonal energy markets used in Texas and California during the sampled time period, are more costly than the bi-settlement nodal markets employed in the Northeast and Midwest, all else equal. Specifications 4 and 5 estimate that implementation of a two-settlement energy market with nodal prices increases costs by twenty-one percent over a single real-time energy market, while the zonal market approach increases costs by forty-six percent. Interestingly, CAISO implemented a two-settlement nodal energy market on April 1, 2009 and ERCOT continues to have ongoing plans to convert their current zonal system to a two settlement nodal design.

5.3. *RTO Effects*

As discussed in section 4, the dummy variables used to control for time-invariant RTO effects can help shed light on the relative efficiency of each RTO. One of the factors—potentially the most important one—subsumed in the RTO fixed effects is efficiency, or ability to minimize costs.[26] To the extent that other RTO-specific unobserved factors are inconsequential, the coefficients on the RTO dummies identify the relative costliness of each RTO.

In the first column of table 5 we find that MISO and SPP appear to be most costly, while the other RTOs are all comparable. However, this specification fails to control for industry-wide changes in costs over time. MISO and SPP were the last two organizations to form. As discussed in section 3, if the model does not control for year-specific cost shocks, and they tend to be increasing, the fixed effects terms of the later forming RTOs will tend to be positively biased. The models that control for time-varying unobserved factors, either with a trend variable or yearly dummies, indicate that all RTOs other than ERCOT have statistically equivalent costs, all else equal. While not a testable conjecture—at least within the scope of this paper—it may be the case that ERCOT's independence from FERC, and the fact that it interacts with a sole state regulator, has conferred a cost advantage over other RTOs.

5.4. *Learning economies*

One of our priors is that costs might fall as RTOs gain experience. This is not confirmed. The coefficient on *LNYRSOP* is positive with a two-sided p-value of 0.054. This indicates that even after controlling for industry wide temporal changes, RTOs become less efficient as they age. One potential explanation is that growing bureaucracy among competing stakeholders adds to costs over time.[27] In column 5, *LNYRSOPTOTAL* is added to explore external learning,

[26] While the notation in section 4 includes a RTO-specific component of the error term, an equivalent way to conceptualize the RTO effect, is to use a RTO-specific technology constant in the underlying production function.

[27] Another possible interpretation is that *LNYRSOP* is correlated with unobserved RTO functions which add to operating expenses. In either case, whether the coefficient on *LNYRSOP* is capturing unobserved determinants of cost, or the model is measuring diseconomies of experience, there is no evidence to support the learning hypothesis.

increased efficiency associated with overall industry experience. Here we find a null effect. While the estimated coefficient on *LNYRSOPTOTAL* is negative, the standard error is roughly three times the magnitude. Taken together, these models provide no evidence of learning, either within or across RTOs.

5.5. *Economies of "scale"*

In the first five models presented in table 5, the estimated coefficients on *LOAD* range from 0.39 to 0.53. All of these estimates provide evidence of increasing returns to scale, as a value of one falls well outside the parameter's ninety-five percent confidence interval in each case. A value of 0.5, which does fall within each of the confidence intervals, implies a production function that is homogenous of degree two. That is, a one hundred percent increase in output is associated with a fifty percent increase in costs.[28]

While increasing returns to scale is not surprising, returns of this magnitude are. It is important to recognize—due to our imperfect measure of RTO scale—the coefficient on *LOAD* arises from two types of variation. PJM and MISO both expanded their geographic footprint during the sample period. Alternatively, the other RTOs experienced growth in output attributed principally to increased demand over a fixed network. When PJM and MISO are excluded from the sample, the estimated elasticity of total costs with respect to output is 0.97, implying constant returns to scale. Thus, the geographic expansion that occurred in PJM and MISO appears to have contributed to the "economies of scale" documented in the other regressions.

It should be noted, if the former is in fact the case, internal learning may in fact occur, however, it remains unidentified by our model.

[28] Our model implies returns to scale are not a function of output. We experimented with models that included a squared term for output; however, the coefficient estimates were unpredictable and erratic. Visual inspection of the residuals (from the models presented) plotted against output, does not indicate a missing variable.

We have attempted to employ other measures of RTO scale, such as miles of transmission wire and the number of transmission-owning members; however, the results have been less than satisfactory. Heterogeneity in the way data are collected across RTOs and the limited number of observations in our data set, both limit our ability to better capture RTO size. Nonetheless, future research should certainly aim to further hash out differences in the economies arising from load growth on a fixed network, from those arising due to geographic expansion.

Our evidence of increasing returns to scale, at least over the sampled range of outputs, is of practical importance. In fact, on July 12, 2001 FERC directed the three Northeast RTOs to enter negotiation to form a single Northeast RTO. The proposal faced strong opposition from stakeholders and state regulators, and eventually faded along with FERC's SMD initiative that had encouraged geographical expansion of the existing "independent transmission providers" (RTOs). On the other hand, in 2004 PJM expanded westward from its original footprint into Ohio, Kentucky, Virginia, West Virginia, Illinois, Indiana and Michigan. PJM and others have argued that this expansion served to lower operating costs and created new opportunities for gains from trade. Political realities and limitations of the existing transmission infrastructure aside, our results provide an additional theoretical argument in favor of larger RTO footprints.

6. Conclusion

Regional transmission organizations are critically important to the restructured electric power sector, but they have drawn increasing scrutiny of their costs, governance, and performance. That scrutiny has not been accompanied by much in the way of systematic evidence or analysis of their performance. This paper has developed a framework for analysis of

RTO costs and has sought to cast light on the major questions raised about their cost experience. Despite some limitations of the data, our analysis comes to several conclusions.

Least surprisingly—despite our inability to recover precise point estimates of the cost of each function—we find that RTO costs are directly related to the number of market functions performed. More specifically, all specifications, including models with RTO and yearly fixed effects, establish that market functions for real-time energy, congestion management (either zonal or nodal transmission rights), and ancillary services each increase costs. In sum, we find convincing evidence that moving beyond their fundamental role as open-access transmission service provider causes RTOs to register notable added costs. Using the parameter estimates from specification 4, we have simulated the estimated total costs of several hypothetical RTOs. The results, presented in table 6, offer a convenient summary of the estimated cost effects of various market functions.

We find no evidence of learning economies, either of an industry-wide or RTO-specific nature. Upon examining economies of scale, we acknowledge that net energy transmitted to load is an imperfect measure of RTO size; however, estimates reveal substantial economies of scale. Analysis that excludes PJM and MISO, which both added transmission owning members, suggests that much of the economies may arise from geographic expansion versus load growth over a fixed network.

Finally, examination of efficiency across RTOs reveals that ERCOT is the only RTO to achieve a different (lower) level of expenditure than PJM, the baseline RTO, after employing a full set of controls. While purely conjecture, one natural interpretation of this result is that ERCOT's less fractured regulatory structure has conferred a cost advantage.

Both our data and model have their limitations, and many have been noted. As a result, we do not claim definitive answers to these questions. Rather, we hope that this approach and its tentative conclusions might prompt further analysis of regional transmission organizations and their role in liberalized electricity markets.

REFERENCES

Barrios, S. and E. Strobl (2004), "Learning by Doing and Spillovers: Evidence from Firm-Level Panel Data from Firm-Level Panel Data." *Review of Industrial Organization*, 25:175-203.

Blumsack, S. (2007), "Measuring the Benefits and Costs of Regional Electric Grid Integration." *Energy Law Journal*, 28:147-184.

Christensen Associated Energy Consulting, LLC. *The Regional Transmission Organization Report Card: Wholesale Electricity Markets and RTO Performance Evaluation Second Edition*, prepared for the National Rural Electric Cooperation Association. August 20, 2007.

Drom R.A. (2007), "New Metrics for Measuring the Success of a Non-Profit RTO." *Energy Law Journal*, 28:603-630.

Lutzenhiser, M. (2004), "Dollars and Sense: Financial Data Raises Doubts about Whether Deregulation Benefits Outweigh Costs." *Public Utilities Fortnightly*, 142:38.

GDS Associates, Inc., *Analysis of Operational and Administrative Costs of RTOs*, prepared for the American Public Power Association. February 5, 2007.

Gruber, H. (1998), "Learning by Doing and Spillovers, Further Evidence from the Semi-Conductor Industry." *Review of Industrial Organization*, 13:697–711.

Irwin, D. A. and P. J. Klenow. (1994), "Learning-by-Doing Spillovers in the Semiconductor Industry." *Journal of Political Economy*, 102:1200–1227.

Joskow, P.L. (2006), "Incentive Regulation in Theory and Practice: Electric Transmission and Distribution Networks." *National Bureau of Economic Research Economic Regulation Project*

Joskow, P.L. (2005a), "Transmission Policy in the United States." *Utilities Policy*, 13:95-115.

Joskow, P.L. (2005b), "The Difficult Transition to Competitive Electricity Markets in the United States." *Electricity Deregulation*. J. Griffin and S. Puller, eds., Chicago, University of Chicago Press.

Public Power Council. *Public Power Council's Response to FERC's "Staff Report on Cost Ranges for the Development and Operation of a Day One Regional Transmission Organization"* November 11, 2004. http://www.ppcpdx.org.

Thornton, R. A. and P. Thompson (2001), "Learning from Experience and Learning from Others: An Exploration of Learning and Spillovers in Wartime Shipbuilding." *American Economic Review*, 91, 1350–1368.

Federal Energy Regulatory Commission. "Staff Report on Cost Ranges for the Development and Operation of a Day One regional Transmission Organization," FERC Docket No. PL04-16-000, October 2004.

U.S. Government Accountability Office (GAO). *Electricity Restructuring: FERC Could Take Additional Steps to Analyze Regional Transmissions Organizations' Benefits and Performance*. GAO-08-987 September 22, 2008.

TABLE 1: RTO Summary Information

	CAISO	ERCOT	ISO-NE	MISO	NYISO	PJM	SPP
First year of data*:	1998	1999	1998	2002	2000	1998	2001
Mean Annual Load (GWh):	229,857	289,101	128,326	553,815	160,487	420,837	194,979
States Covered:	California	Texas	Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island and Vermont	North Dakota, South Dakota, Nebraska, Minnesota, Iowa, Wisconsin, Illinois, Indiana, Michigan and parts of Montana, Missouri, Kentucky, and Ohio.	New York	Pennsylvania, New Jersey, Maryland, Delaware, District of Columbia, Virginia, West Virginia and Ohio. Parts of Illinois, Michigan, Indiana, Kentucky, North Carolina and Tennessee.	Kansas, Oklahoma and parts of New Mexico, Texas, Louisiana, Missouri, Mississippi and Arkansas
Headquarters:	Folsom, California	Austin, Texas	Holyoke, Massachusetts	Carmel, Indiana	Rensselaer, New York	Valley Forge, Pennsylvania	Little Rock, Arkansas
Installed Generation (MW):	54,500	71,812	32,000	156,000	44,851	164,634	50,392
Miles of Transmission Lines:	25,526	38,000	8,000	93,600	11,000	56,250	40,364
Population Served (millions):	30	20	14	40	19	51	5

*The first annual data record for each RTO corresponds to the year which it began providing regional transmission service under a FERC open access transmission tariff. Two exceptions follow: (1) ERCOT is not under FERC jurisdiction, and hence, does not file a tariff with them. On August 21, 1996, the Public Utility Commission of Texas required ERCOT to form an ISO with the responsibilities of grid security, market facilitation, and coordination of transmission planning. On May 27, 1999, the Texas legislature passed S.B. No. 7, which provides a framework for open access transmission. In addition to operating a control area and its traditional function of preserving reliability, ERCOT was charged with ensuring access to transmission services and settling accounts. (2) SPP commenced open access transmission operations for service across 11 transmission owners in 1998; however, 2001 was the first year of publically available financial statements.

TABLE 2: Implementation Dates of Market Functions

	CAISO	**ERCOT**	**ISO-NE**	**MISO**	**NYISO**	**PJM**	**SPP**
Real-time energy	4/98	8/01	5/99	4/05	12/99	4/98	2/07
Day-ahead energy			3/03	4/05	12/99	6/00	
FTR	2/00	2/02	3/03	4/05	12/99	5/99	
Ancillary Services[29]	4/98	8/01	5/99		12/99	Regulation 6/00 Reserve 12/02	
Capacity Markets			5/99 – 8/00 Reopened 4/03		12/99	1/99	

[29] *ANCILL* equals one for PJM in 2001 and later years.

TABLE 3: List of Variables

Variable	Definition
OPEXP	Log of total operating expenses, including operations, maintenance, and depreciation expenses (2005 dollars)
REALTIME	Dummy = 1 for real-time energy market
DAYAHEAD_FTR	Dummy = 1 for day-ahead energy and nodal transmission rights
ZONE_FTR	Dummy = 1 for zonal transmission rights
ANCILL	Dummy = 1 for ancillary services market
CAPACITY	Dummy = 1 for capacity market
LNYRSOP	Log of the RTO's age in year *t*.
LNYRSOPTOTAL	Log of the sum of all RTOs' ages in year *t*.
WAGE	Log of average annual pay for the electric power generation, transmission, and distribution sector, NAICS 2211 (2005 dollars)
LOAD	Log of net energy for load (GWh)

TABLE 4: Major FERC Orders Related to RTOs

Order No. 719: Wholesale Competition in Regions with Organized Electric Markets Issued October 17, 2008 FERC amends regulations to improve operation of RTO markets in the areas of: (1) demand response; (2) long-term power contracting; (3) market-monitoring policies; and (4) the responsiveness of RTOs to their customers and other stakeholders. Each RTO is required to make filings that ensure its tariff and market design satisfies requirements in each area.
Order No. 668: Accounting and Financial Reporting for Public Utilities Including RTOs Issued December 16, 2005 FERC amends its regulations on accounting rules for RTOs and requires new quarterly financial reporting.
Policy Statement on Market Monitoring Units (MMUs) Issued May 27, 2005 FERC sets out five protocols for referrals by an RTOs MMUs to the Commission of: (1) alleged tariff violations; and (2) alleged violations of Market Behavior Rules.
Policy Statement on Credit-Related Issues Issued November 19, 2004 FERC asks RTOs to: (1) make their credit-related procedures and standards more transparent; and (2) better assess transmission customers' creditworthiness. The order also clarified certain actions that FERC expects RTOs to take to reduce the risk and impact of a default by a market participant.
Order No. 2000: Regional Transmission Organizations Issued December 20, 1999 FERC amends its regulations to advance the formation of RTOs and codifies minimum characteristics and functions that a transmission entity must satisfy in order to be considered an RTO.

Source: The Federal Energy Regulatory Commission, http://www.ferc.gov/industries/electric/indus-act/rto/maj-ord.asp

TABLE 5: Regression Parameter Estimates

	(1) OPEXP	(2) OPEXP	(3) OPEXP	(4) OPEXP	(5) OPEXP	(6) OPEXP
REALTIME	0.371**	0.362**	0.566**	0.437**	0.437**	0.340**
	(0.10)	(0.09)	(0.08)	(0.10)	(0.10)	(0.14)
DAYAHEAD_FTR	0.407**	0.243**	0.165*	0.194*	0.194*	0.188
	(0.10)	(0.10)	(0.09)	(0.10)	(0.10)	(0.14)
ZONE_FTR	0.351*	0.436**	0.534**	0.376**	0.376**	0.365**
	(0.18)	(0.17)	(0.17)	(0.17)	(0.17)	(0.18)
ANCILL	0.356**	0.363**	0.101	0.164	0.164	0.562**
	(0.11)	(0.11)	(0.11)	(0.13)	(0.13)	(0.23)
CAPACITY	-0.133	-0.139	-0.007	0.033	0.033	
	(0.12)	(0.12)	(0.13)	(0.14)	(0.14)	
LOAD	0.530**	0.389**	0.500**	0.487**	0.487**	0.975
	(0.13)	(0.14)	(0.10)	(0.11)	(0.11)	(0.64)
WAGE	2.062**	0.442	0.05	0.7	0.7	-0.072
	(0.69)	(0.76)	(0.65)	(0.73)	(0.73)	(0.83)
CAISO	0.197	-0.099	0.081	0.214	0.214	-0.076
	(0.27)	(0.37)	(0.26)	(0.23)	(0.23)	(0.49)
ERCOT	-0.32	-0.764**	-0.724**	-0.452**	-0.452**	-0.917
	(0.22)	(0.22)	(0.21)	(0.23)	(0.23)	(0.57)
ISONE	-0.048	-0.231	-0.036	-0.062	-0.062	
	(0.18)	(0.18)	(0.13)	(0.16)	(0.16)	
MISO	0.771**	0.283	0.078	0.497*	0.497*	
	(0.18)	(0.22)	(0.21)	(0.29)	(0.29)	
NYISO	0.094	-0.118	0.046	0.149	0.149	-0.086
	(0.16)	(0.16)	(0.10)	(0.12)	(0.12)	(0.20)
SPP	0.440**	-0.323	-0.382*	-0.028	-0.028	-0.25
	(0.17)	(0.25)	(0.21)	(0.27)	(0.27)	(0.40)
YEAR		0.069**				0.074**
		(0.02)				(0.02)
LNYRSOP				0.218*	0.218*	
				(0.11)	(0.11)	
LNYRSOPTOTAL					-0.042	
					(0.13)	
Observations	67	67	67	67	67	49
Year fixed effects	No	No	Yes	Yes	Yes	No
Wald Chi-squared	738.29	932.29	2472.06	1495.46	1495.46	323.74

Standard errors in parentheses

* significant at 10%; ** significant at 5%.

TABLE 6: Simulation of Total Costs*

Hypothetical RTO	Total Costs	Change in costs			
Case (A) = Open-access transmission service provider	$68,418,713				
Case (B) = Case (A) + Real-time energy	$105,864,913	$37,446,199	or	55%	greater costs than case (A).
Case (C) = Case (B) + Ancillary services	$124,754,558	$18,889,646	or	18%	greater costs than case (B).
Case (D) = Case (C) + Zonal transmission rights	$181,740,760	$26,637,680	or	21%	greater costs than case (C).
Case (E) = Case (C) + Day-ahead and nodal FTRs	$151,392,238	$56,986,201	or	46%	greater costs than case (C).

*Predicted values arise from parameter estimates of model (4) in table 5. An average of the RTO fixed effects is used along with the year 2008 fixed effect and the mean value of *WAGE*, *LOAD*, and *LNYRSOP* in 2008.